REMARKS/ARGUMENTS:

This paper is filed in response to the Examiner's Office Action dated October 12, 2006 for the above-captioned U.S. Patent Application. The office action objected to the specification and rejected claims 1-46.

The applicant would respectfully advise the Examiner that another Information Disclosure Statement is herewith provided including full citation to references cited in the prior IDS submitted on June 7, 2004. The Examiner is requested to also consider these references and to make them of record.

The Examiner has rejected claims 1-46 under 35 USC 112, second paragraph, as indefinite, and has made the following further rejections:

claims 1-4, 7, 9-15, 17, 19-22, 25, 27-33, 35 and 37-40 under 35 USC 102(b) as anticipated by Garfield;

claims 18 and 36 under 35 USC 103(a) as obvious over Garfield and Kemp;

claims 5, 16, 23, 34 and 41 under 35 USC 103(a) as obvious over Garfield and Dittmar;

claims 6, 24 and 42 under 35 USC 103(a) as obvious over Garfield and Leiter;

claims 8 and 26 under 35 USC 103(a) as obvious over Garfield and Drefahl;

claim 43 under 35 USC 103(a) as obvious over Garfield and Shivaratri;

claim 44 under 35 USC 103(a) as obvious over Garfield, Shivaratri, and Leiter; and

claims 45 and 46 under 35 USC 103(a) as obvious over Garfield, Shivaratri, and Drefahl.

Claims 1-3, 6, 19-21, 24, 37-39 and 42-44 are herein amended, support for which is found at least at page 12 lines 8-13; page 13 line 20 to page 14 line 20; and page 20 lines 1-5. No new matter is added.

Claim 43 has been amended as suggested by the Examiner, to address the rejection under 35 USC 112, 2nd paragraph. Withdrawal of that rejection to claim 43, as well as to dependent claims 44-46, is respectfully requested.

16

The rejection under 35 USC 112, 2nd paragraph to claims 1, 19 and 37 is seen to be in error. The office action recites "There is no indication that a component in a biological sample is to be identified as resuscitated in the preamble". This subject matter does not appear to be related to claims 1, 19 or 37, or to this application. Elements of claims 1, 19 and 37 are seen to describe processing a document as recited in each claim's preamble. The final element "storing" is not seen to negate that processing specifically done in other claim elements (e.g. partitioning). The applicant respectfully requests the Examiner withdraw the rejection under 35 USC 112, 2nd paragraph of claims 1, 19 and 37, as well as claims depending from them. In the alternative, the Applicants request the Examiner issue a new non-final office action that clarifies the rejection.

The invention relates to a system and a method to identify organic chemical nomenclature from text documents, and from that information to index chemical fragments and their structures and connectivity (page 7, line 4-6). In the indexing phase, chemically related documents such as chemical and drug patents or articles are scanned and indexed (page 12 line 21 to page 13 line 1). A text index is also made that includes keywords (e.g., antidepressant) and a document identifier (e.g., title, author), so that a searcher may search a database generated as above by text and by chemical structure/substructure (page 20 lines 1-5; page 14, lines 14-18).

Garfield relates to an "algorithm for translating chemical nomenclature into molecular formulas," (title, page 445, second para, first sentence; page 460, second full para, last sentence; page 460, third full para, third sentence).

Claim 1, taken as exemplary of the distinctions of the independent claims over the art, recites:

> **A method to process a document, comprising: partitioning document text and assigning semantic meaning to words of the partitioned text, where assigning comprises applying a plurality of regular expressions, rules and a plurality of dictionaries to recognize chemical name fragments; recognizing any substructures present in the chemical name fragments; determining structural connectivity information of the**

17

> **chemical name fragments and recognized substructures; extracting identifying information from the document; and storing the extracted identifying information in association with the determined structural connectivity information in a searchable index.**

Garfield discloses that the "the entire chemical name is punched on an IBM card or typed directly on a Unityper typewriter," (page 485, second full para). Further, Garfield discloses that "working from right to left each character in the name is brought into the computer register one at a time and processed one at a time," and "the character in process at any instant is referred to as the current character," (page 485, second full para). Further, in Garfield, "each character (of the chemical name) is processed." Garfield reads characters and groups of characters from the punch card until a dictionary match is found (page 486-488). The Examiner apparently considers Garfield's punch card as anticipating the claimed document. Because Garfield neither extracts identifying information from its punch cards, nor stores extracted punch card identifying information in association with anything else in a searchable index, claim 1 is seen to be novel over Garfield.

Garfield is directed to translating chemical names to chemical formulas, whereas the present invention is directed to providing a searchable database. Both deal with the problem of different nomenclature used for the same underlying chemical structure, but whereas the goal of Garfield is to produce a structural diagram from an input chemical name (page 460, second full para, last sentence; page 460, third full para, third sentence), one goal of the present invention is a searchable database. Thus Garfield's punch card inputs are isolated from any document that may include an identifier (or keywords as in dependent claims 2-3 and 6). Claim 1 as amended extracts identifying information from the same document from which the chemical name fragments are recognized. By associating these with one another in the database, a searcher, using a chemical name fragment or fragment structure as a search term, can find a list of relevant documents without having the specify the actual chemical name explicitly recited in those documents. (see page 10 line 22 to page 11 line 2; page 13 lines 3-8; page 14 lines 13-20). Garfield lacks both document identifying information and a stored association of that identifying information with structural connectivity information as claim 1 recites.

18

While not required, below is shown that claim 1 is further non-obvious over ordinary skill modifications to Garfield. Garfield's principle of operation is to generate a structural diagram from an input chemical name. As is clear from the citations above, 1) the database of this application is generated from underlying documents so that an entire chemical name need not be entered to search the database, whereas Garfield uses the full chemical name in every instance of his punchcards; and 2) identifying documents that match a searcher's entered chemical structure (and in dependent claims also an entered keyword) is a goal of setting up the database in the manner described in this application, whereas Garfield's chemical names are entered via punchcard or typewriter and not as part of a document that might also include identifying information (and keywords). In fact, Garfield has no use for such identifying information or keywords since those do not add to his goal of drawing a chemical structure, which Garfield does from entering only the chemical name. Neither Garfield nor any other reference is seen to associate document identifying information with structural connectivity information of chemical name fragments that are recognized in that same document, as recited in claim 1. Claim 1 is therefore seen to be non-obvious over Garfield in combination with any other art of record.

Specifically, Leiter's registry is limited to an association of chemical structure to chemical name for new compounds entered into the CAS registry, and is not seen to associate connectivity information of chemical name fragments that are recognized in a document with identifying information extracted from that document. Kemp parses text of underlying documents (e.g., patents) to find names of non-generic chemical substances and to differentiate chemical from non-chemical names (page 545, lefthand column, second full paragraph; page 547, lefthand column, first full paragraph). Kemp is not seen to utilize the non-chemical names. As to the chemical names, Kemp's approach appears limited to the explicit text recited in the underlying document, merely accounting for different punctuations (e.g., dash, space, hyphen) that may or may not be used by different document authors to separate chemical name fragments within a singular chemical name. See Kemp's example at page 548, lefthand column, first full paragraph and the referenced Table 5. There is seen no structural connectivity information in Kemp, only text-matching that

19

resolves different punctuations (or lack thereof) between chemical fragment names of a compound.

For reasons analogous to those detailed above with respect to claim 1, similarly amended independent claims 19, 37 and 43 are seen to be novel over Garfield as well as non-obvious in view of any combination of cited art. The above distinctions are seen to be substantial, and so the dependent claims are not separately argued without conceding the Examiner's rejection of any, with the few exceptions below.

Claims 2-3 are amended to follow the amendment to independent claim 1 from which they depend, and further to include extracting keywords from the document and searching the index by a keyword and at least one other term. Dependent claims 20-21 and 38-39 are amended similarly to claims 2-3.

Claim 6 is also amended with subject matter related to keywords, and specifically recites that extracted keywords are stored in a text index whereas the structural connectivity is stored in a structure index. Dependent claims 24, 42 and 44 are amended similarly to claim 6.
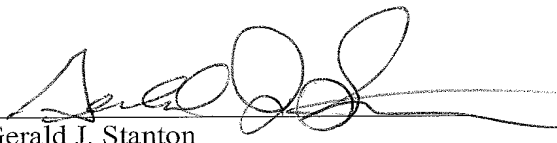
Respecting keywords, no other reference is seen to extract keywords from a same document from which chemical name fragments are recognized. The cited art to Garfield, Leiter and Kemp recognize that different chemical names or nomenclatures can be used to refer to an identical chemical structure. Each therefore teaches away from searching a database (that may be generated by their teachings) using nomenclature or chemical name, so a chemical name as keyword cannot anticipate or render obvious the keywords of the above dependent claims.

The Applicants thank the Examiner for a thorough search. In view of the above amendments and remarks, claims 1-46 are seen to be in condition for allowance and the Applicants respectfully request the Examiner pass them to issue. The undersigned

20

representative welcomes the opportunity to resolve any matters that may remain, formal or otherwise, via teleconference at the Examiner's discretion.

Respectfully submitted:

_____   _____January 12, 2006_____

Gerald J. Stanton   Date
Reg. No.: 46,008

Customer No.: 48237
HARRINGTON & SMITH, LLP
4 Research Drive
Shelton, CT 06484-6212
Phone: (203) 925-9400
Facsimile: (203) 944-0245
Email: gstanton@hspatent.com